



Centre d'Estudis Demogràfics

**EL ESTUDIO DE LA MIGRACIÓN INTERNACIONAL
LATINOAMERICANA A PARTIR DE
LA BASE DE MICRODATOS CENSALES
DE IPUMS INTERNATIONAL**

Antonio LÓPEZ-GAY
Robert McCAA
Albert ESTEVE

380

*PAPERS
DE
DEMOGRAFIA*

2011



Centre d'Estudis Demogràfics

**EL ESTUDIO DE LA MIGRACIÓN INTERNACIONAL
LATINOAMERICANA A PARTIR DE
LA BASE DE MICRODATOS CENSALES
DE IPUMS INTERNATIONAL**

Antonio LÓPEZ-GAY
Robert McCAA
Albert ESTEVE

380

Centre d'Estudis Demogràfics

2011

Resum.- *L'estudi de la migració internacional llatinoamericana a partir de la base de microdades censals de IPUMS International*

Les microdades censals són una de les fonts d'informació fonamentals per a l'estudi sociodemogràfic de la migració internacional. El projecte *IPUMS-International* s'ha encarregat, des de 1999, de recopilar censos de població de tot el món, i d'homogeneïtzar la seva informació (tant les microdades com les metadades), amb l'objectiu de proporcionar a l'investigador, dades comparables a nivell temporal i espacial. Actualment, el sistema de recerca d'IPUMSi, a través d'Internet, inclou 35 països a on, els d'Amèrica Llatina -la seva participació ha liderat el projecte mundial- constitueix gairebé la tercera part d'aquests països. Dins d'Amèrica Llatina, Mèxic va ser pioner en el projecte gràcies al suport constant de l'INEGI, que va aportar quatre mostres censals d'alta precisió, a més de dos recomptes. Tot això converteix a la base de *IPUMS-International* en una de les eines més valuoses per a l'anàlisi del fenomen migratori mundial, especialment en el context llatinoamericà. Aquest treball doncs, descriu les possibilitats i limitacions d'aquesta base de dades, emfatitzant en la disponibilitat de les variables més interessants per a l'estudi del fet migratori.

Paraules clau.- Migració internacional, IPUMS-International, microdades censals, Amèrica Llatina.

Resumen.- *El estudio de la migración internacional latinoamericana a partir de la base de microdatos censales de IPUMS International*

Los microdatos censales son una de las fuentes de información fundamentales para el estudio sociodemográfico de la migración internacional. El proyecto *IPUMS-International* se ha encargado, desde 1999, de recopilar censos de población de todo el mundo, y de homogeneizar su información (tanto los microdatos como los metadatos), con el objetivo de proporcionar al investigador datos comparables a nivel temporal y espacial. En la actualidad, el sistema de extracción de IPUMSi a través de Internet, incluye 35 países, y América Latina, cuya participación ha liderado el proyecto mundial, constituye casi la tercera parte de ellos. En América Latina, México fue pionero en el proyecto gracias al apoyo constate del INEGI, que aportó cuatro muestras censales de alta precisión además de dos conteos. Todo esto convierte a la base de *IPUMS-International* en una de las herramientas más valiosas para el análisis del fenómeno migratorio mundial, especialmente en el contexto latinoamericano. El documento describe las posibilidades y limitaciones de esta base de datos, enfatizando en la disponibilidad de las variables más interesantes para el estudio del hecho migratorio.

Palabras clave.- Migración internacional, IPUMS-International, microdatos censales, América Latina.

Abstract.- *The study of international migration from Latin American based International IPUMS census microdata*

Census microdata constitute one of the fundamental sources for the social and demographic study of International migration. Since 1999, the IPUMS-International initiative has sought to recover microdata world-wide, and integrate anonymized samples as well as metadata, and make them available to researchers in a form suitable for comparative research in space and time. Currently, the IPUMS extraction system includes data for 35 countries via the internet. Latin America, whose participation has been that of a leader in the project, constitutes almost one-third of the total. Within Latin America, Mexico, through the constant support of INEGI, pioneered the project, entrusting four high precision census samples plus two samples of population counts (“conteos”). All this converts the IPUMS-International data base in one of the most valuable tools for the analysis of the global migration phenomenon, especially within the Latin American context. The present paper is a methodological offering which discusses the strengths and weaknesses of the IPUMS database, with an emphasis on the availability of variables of greatest interest for the study of migration.

Keywords.- International migration, IPUMS-International, census microdata, Latin America.

ÍNDICE

1.- El proyecto IPUMS-International: antecedentes y objetivos	1
1.1.- Antecedentes	1
2.2.- Objetivos del proyecto IPUMS-International: preservar, integrar y difundir	5
2.- Comparabilidad espacial y temporal de los microdatos censales	6
2.1.- Métodos en la integración internacional de microdatos	6
2.2.- El sistema de extracción de microdatos censales	8
3.- El estudio de la migración internacional a partir de la base de datos IPUMS-International	9
3.1.- Muestras censales disponibles para el estudio de la migración internacional en el contexto latinoamericano	9
3.2.- Disponibilidad de variables para el estudio de la migración internacional en IPUMSI	15
3.3.- Comparabilidad espacial y temporal de la información: nuevos enfoques en el estudio de la migración internacional	21
4.- A modo de recapitulación	23

ÍNDICE DE CUADROS

1.- Censos de población elaborados desde 1960 en América Latina y disponibilidad en el sistema IPUMSI	3
2.- Características de las muestras censales de América Latina disponibles en el sistema IPUMSI	4
3.- Población nacida en el extranjero y listado de los países de nacimiento con mayor presencia en los censos de América Latina disponibles en la base de datos IPUMS-International	10
4.- Población nacida en el extranjero y listado de los países latinoamericanos de nacimiento en los censos de los países del mundo con mayor presencia de población de América Latina (disponibles en IPUMSI)	13
5.- Disponibilidad de variables migratorias y de naturaleza en los censos de los países de América Latina (disponibles en IPUMSI), España y Estados Unidos ..	16
6.- Disponibilidad de variables de situación migratoria en los censos de los países de América Latina (disponibles en IPUMSI), España y Estados Unidos	18
7.- Ejemplo de codificación en el sistema IPUMSI de la variable “Situación migratoria, residencia anterior”	19
8.- Conjunto de variables del hogar disponibles en la base de datos IPUMSI	20
9.- Conjunto de variables del persona disponibles en la base de datos IPUMSI	21

**EL ESTUDIO DE LA MIGRACIÓN INTERNACIONAL LATINOAMERICANA
A PARTIR DE LA BASE DE MICRODATOS CENSALES
DE IPUMS INTERNATIONAL**

Antonio LÓPEZ-GAY
tlopez@ced.uab.es

Robert McCAA
rmccaa@umn.edu

Albert ESTEVE
aesteve@ced.uab.es

1.- El proyecto IPUMS-International: antecedentes y objetivos

1.1.- Antecedentes

La idea de proporcionar a la comunidad científica microdatos censales integrados se desarrolla a lo largo de la década de 1990 con la puesta en marcha del proyecto IPUMS-USA (Integrated Public Use Microdata Series- USA) desde el Minnesota Population Center de la Universidad de Minnesota. El proyecto puso a disposición gratuita de los investigadores las series de microdatos armonizados pertenecientes a los censos que se llevaron a cabo en Estados Unidos entre 1850 y 2000. Desde el inicio de esta iniciativa, la base de datos IPUMS-USA se consolidó como una de las fuentes de información demográfica más utilizadas entre los investigadores en ciencias sociales.

En 1998, por primera vez, se extendió el paradigma de IPUMS a los censos de Colombia, en una experiencia piloto realizada en estrecha colaboración con el *Departamento Nacional de Estadística de Colombia* (DANE). Ésta fue la primera la primera piedra de lo que posteriormente ha sido una exitosa iniciativa de integración de microdatos censales a escala internacional: IPUMS-International (IPUMSI). En 1999, la agencia estadounidense *National Science Foundation* financió la propuesta IPUMSI, acontecimiento fundamental para la construcción de lo que hoy en día es la base de microdatos censales más grande del mundo.

En la actualidad, IPUMSI pone al alcance de los investigadores de forma gratuita los microdatos correspondientes a 111 muestras censales de 35 países de todo el mundo. En total, la base de datos suma más de 250 millones de registros de personas, una cifra que se amplía año tras año ya que son más de 80 los institutos de estadística que han firmado un convenio de colaboración con la Universidad de Minnesota para la difusión de sus microdatos censales (Cuadro 1).

La colaboración de los institutos de estadística de América Latina se ha configurado como pieza clave en el positivo desarrollo del proyecto global. Gracias a su espíritu científico, la región ya es líder en el proyecto. Los microdatos armonizados de 43 muestras censales correspondientes a 9 países latinoamericanos están a la disposición de los más de 2.000 investigadores de todo el mundo que se han registrado en el sistema de extracción de datos del proyecto IPUMSI. Aunque la proporción total de habitantes que residen en América Latina no alcanza la décima parte de la población mundial, su representación en el IPUMSI supera la tercera parte del banco de datos, de forma que los microdatos del conjunto de países América Latina suman más de 100 millones de registros anónimos (Cuadro 2).

Cuadro 1.- Censos de población elaborados desde 1960 en América Latina, y disponibilidad en el sistema IPUMSI

La totalidad de muestras disponibles actualmente se pueden consultar en www.ipums.org/international

País	Muestras	Notas	Década censal				
			2000s	1990s	1980s	1970s	1960s
Muestras anonimizadas y homologadas disponibles en la base de datos IPUMS							
Argentina	4	Completo	2001	1991	1980	1970	1960
Brasil	5	Completo	2001	1991	1980	1970	1960
Colombia	5	Completo	2005	1993	1985	1973	1964
Costa Rica	4	Completo	2000		1984	1973	1963
Chile	5	Completo	2002	1992	1982	1970	1960
Ecuador	5	Completo	2001	1990	1982	1974	1962
México	6	Completo	2000-05	1990-95	1980	1970	1960
Panamá	5	Completo	2000	1990	1980	1970	1960
Venezuela	4	Completo	2001	1990	1981	1971	1961
Muestras de otros países del mundo de interés para el estudio de la migración internacional latinoamericana							
Canadá	4	Completo	2001	1991-6	1981-6	1971-6	1961-6
España	3	Completo	2001	1991	1981		
Estados Unidos	6	Completo	2000-05	1990	1980	1970	1960
Portugal	3	Completo	2001	1991	1981		
Muestras para inaugurar en años próximos durante el IPUMS-AL II							
El Salvador	1		2007	1992		1971	1961
Guatemala	5		2002	1994	1981	1973	1964
Honduras	4		2000		1988	1974	1961
Nicaragua	1		2005	1995		1971	1963
Puerto Rico	4		2000	1990	1980	1970	1960
Saint Lucia	2		2001	1991	1980	1970	1960
Bolivia	3		2001	1992		1976	
Paraguay	5		2002	1992	1982	1972	1962
Perú	1		¿?	1993	1981?	1972	1961
Rep. Dominicana	4		2002	1993	1981	1970	1960
Uruguay	4			1996	1985	1975	1963

Nota: año en **negrita** = existencia de microdatos confirmados. Para precisar atrasos en caso específicos, favor de consultar con el delegado de IPUMS o enviar email a rmccaa@umn.edu

Cuadro 2.- Características de las muestras censales de América Latina disponibles en el sistema IPUMSI

Censo	Tamaño de la muestra (%)	Número de Hogares	Número de personas
Argentina 1970	2	129.728	466.892
Argentina 1980	10	672.062	2.667.714
Argentina 1991	10	1.148.351	4.143.727
Argentina 2001	10	1.040.852	3.626.103
Brasil 1960	5	613.273	3.001.439
Brasil 1970	5	1.022.207	4.953.759
Brasil 1980	5	1.343.377	5.870.467
Brasil 1991	5,8	2.012.276	8.522.740
Brasil 2000	6	2.652.356	10.136.022
Chile 1960	1	n.d	88.184
Chile 1970	10	199.041	890.481
Chile 1982	10	282.356	1.133.062
Chile 1992	10	373.964	1.335.055
Chile 2002	10	486.115	1.513.914
Colombia 1964	2	n.d	349.652
Colombia 1973	10	349.859	1.988.831
Colombia 1985	10	571.046	2.643.125
Colombia 1993	10	774.321	3.213.657
Colombia 2005	10	1.054.901	4.117.607
Costa Rica 1963	5	n.d.	82.345
Costa Rica 1973	10	36.323	186.762
Costa Rica 1984	10	56.186	241.220
Costa Rica 2000	10	106.973	381.500
Ecuador 1962	3	n.d	136.443
Ecuador 1974	10	145.902	646.678
Ecuador 1982	10	195.401	806.834
Ecuador 1990	10	243.898	966.234
Ecuador 2001	10	354.222	1.213.725
México 1960	1.5	n.d	502.800
México 1970	1	82.856	483.405
México 1990	10	1.648.280	8.118.242
México 1995	0.4	72.277	332.061
México 2000	10.6	2.312.035	10.099.182
México 2005	10	2.546.985	10.284.550
Panamá 1960	5	11.869	53.553
Panamá 1970	10	31.755	150.473
Panamá 1980	10	47.726	195.577
Panamá 1990	10	61.458	232.737
Panamá 2000	10	84.346	284.081
Venezuela 1971	10	284.336	1.158.527
Venezuela 1981	10	323.321	1.441.266
Venezuela 1990	10	468.808	1.803.953
Venezuela 2001	10	646.080	2.306.489

Fuente: <https://international.ipums.org>

1.2.- Objetivos del proyecto IPUMS-International: preservar, integrar y difundir

Los tres grandes pilares sobre los que se sustenta el proyecto IPUMSI son los de la preservación, la integración y la difusión de los microdatos censales.

a) *Preservar*: Uno de los principales objetivos que se planteó el proyecto IPUMSI desde sus inicios fue el de inventariar los microdatos y documentos censales que han sobrevivido hasta nuestros días. Este conjunto de información es de extrema importancia para la posterior integración de los microdatos censales en la base de datos IPUMSI, además de representar un valioso recurso para los investigadores y archiveros de datos censales. El proyecto ha conseguido recopilar la gran mayoría de cuestionarios censales y la documentación adicional de los censos de población que se han realizado en todo el mundo durante los últimos 50 años.¹

b) *Integrar*²: La integración de los microdatos censales es el principal reto al que IPUMSI debe enfrentarse, pero también su principal fortaleza, el hecho diferencial que singulariza a esta base de datos respecto a todas las demás. Sin lugar a dudas, el inventariado previo de la documentación censal es fundamental para conocer las definiciones y conceptos que cada país ha utilizado para la realización de su censo de población. Durante la etapa de integración, los datos son procesados con cuatro finalidades distintas: i) garantizar la confidencialidad de los datos; ii) reformar, limpiar e imputar valores perdidos en la base de datos; iii) armonizar variables; iv) construir variables. IPUMSI también elabora documentación homogeneizada para cada una de las muestras y variables con el objetivo de proporcionar al usuario final información relevante del proceso de integración.

c) *Difundir*: La difusión de los más de 250 millones de registros anónimos que actualmente forman parte de la base de datos IPUMSI requiere una plataforma eficiente.

¹ Disponibles en breve en: <https://international.ipums.org>

Ésta debe permitir hacer frente no sólo a las limitaciones que representa el trabajar con una base de datos de enorme tamaño, sino también debe proporcionar al usuario una clara y sencilla forma de acceder a la información. La difusión se realiza a través de Internet (sitio <https://international.ipums.org/international>), mediante un sistema gratuito que permite al usuario confeccionar su propia base de datos, escogiendo formatos, muestras, variables y casos específicos. Además, se ha potenciado la posibilidad de obtener información sobre todas las variables de forma sencilla, de manera que el usuario puede aclarar las dudas que eventualmente surjan a lo largo del proceso de extracción de los microdatos.

2.- Comparabilidad espacial y temporal de los microdatos censales

2.1.- Métodos en la integración internacional de microdatos

El proceso de homologación de microdatos censales requiere del estudio minucioso de los conceptos y definiciones de cada una de las cuestiones en el formulario censal, de forma que el proyecto pretende homologar, simultáneamente, no sólo las bases de datos sino la documentación de los censos tanto en el tiempo como en el espacio. Han cambiado mucho los conceptos censales desde los años sesenta hasta la actualidad y cada país tiene sus prácticas y convenciones propias. Además, las muestras originales varían significativamente en cuanto a la calidad, formatos, conceptos tratados, y esquemas de clasificación de variables. A continuación se exponen las diferentes técnicas y procedimientos que IPUMSI afronta a la hora de armonizar los microdatos censales:

a) ***Garantizar la confidencialidad de los datos.*** IPUMSI ha diseñado dos estrategias para alcanzar dicho objetivo. Primero, a través del vínculo con el investigador, que se compromete a preservar la confidencialidad de los datos con la firma de un protocolo en el momento en el que realiza el registro como usuario. Segundo, a través de medidas de protección estadística, tales como la modificación aleatoria del orden de los hogares, la

² A lo largo del texto, también se define este proceso como armonización y homologación.

combinación de códigos que informan de características sensibles o la agrupación de las entidades geográficas que no superan los 20.000 habitantes.

- b) *Estandarizar formatos, limpiar e imputar valores perdidos de la base de datos.*** La gran mayoría de las muestras censales proporcionadas por los institutos de estadística internacionales permiten adaptarlas a un tipo jerárquico en el que se combina información referente a los hogares y a las personas. Además, IPUMSI examina la consistencia interna de la base de datos e imputa, cuando es posible, los valores perdidos.
- c) *Integrar variables.*** La integración de variables ocupa un lugar central en el proceso de integración. El objetivo que se persigue en esta fase es la creación de variables coherentes en el tiempo y el espacio que permitan la comparación entre países y en perspectiva histórica. La heterogénea concepción de los cuestionarios censales, así como la diversidad en la formulación de las preguntas y respuestas, hacen que el proceso de armonización sea complejo por propia definición. Por estas razones, IPUMSI identifica y documenta todas las incompatibilidades potenciales. El objetivo final en el proceso de integración es el de crear nuevas variables a partir de un ejercicio de recodificación, que debe satisfacer dos requisitos: garantizar la máxima comparabilidad en el tiempo y en el espacio y, simultáneamente, retener todo el detalle contenido en las variables originales. Para ello, IPUMSI utiliza un sistema de codificación compuesto de varios dígitos, todos ellos armonizados.
- d) *Construir variables.*** IPUMSI construye nuevas variables para maximizar la utilización de los datos ya armonizados. Se han creado variables a nivel individual que informan de las relaciones entre miembros de la familia, facilitando así la tarea a los investigadores que quieran identificar hogares con una determinada estructura. Para cada individuo, se construyen tres variables que identifican la posición de la madre, el padre y la pareja (en matrimonio o unión libre) dentro del hogar.

2.2.- El sistema de extracción de microdatos censales

El acceso a la documentación y a los metadatos es totalmente accesible a través de la página web, sin ninguna restricción. En cambio, como se ha comentado anteriormente, los usuarios deben solicitar acceso para poder extraer los microdatos censales, un proceso sencillo y de rápida confirmación que pretende asegurar el buen uso de los datos por parte del usuario.

El sistema de extracción de microdatos que el investigador debe utilizar cada vez que desea realizar una petición de datos es claro y conciso y se estructura en seis grandes pasos. Cada una de estas etapas conduce al usuario a la selección de los parámetros que definirán su extracto final. El investigador debe personalizar su petición seleccionando las muestras censales de las que desea obtener datos, así como las variables que necesita para su estudio. Es importante destacar que el sistema de solicitud de datos también permite incluir las variables con sus respectivas categorías no armonizadas, otro elemento más que proporciona al investigador un mayor control sobre todo el proceso de integración que el equipo IPUMSI ha llevado a cabo. Posteriormente, el usuario también puede asociar características de otros miembros del hogar a ese individuo, así como seleccionar los registros personales o del hogar que cumplen una determinada condición. Una reciente novedad del sistema ha sido la inclusión de una aplicación que permite personalizar el tamaño de la muestra, siempre y cuando el investigador necesite una segunda muestra más pequeña que la que proporciona la base de datos IPUMSI.

Una vez completado el proceso de petición de datos, y pasados unos minutos, el usuario recibe un correo electrónico en el que se suministra el vínculo que conduce a la página de descargas personales. Ésta incluye el archivo de microdatos censales solicitado. Además, también se proporciona al usuario un conjunto de archivos de sintaxis que le permiten importar los microdatos de forma directa a cualquiera de los paquetes estadísticos más utilizados (SPSS, SAS y STATA).

3.- El estudio de la migración internacional a partir de la base de datos IPUMS–International

3.1.- Muestras censales disponibles para el estudio de la migración internacional en el contexto latinoamericano

Los censos de población son una de las fuentes de información fundamentales para el estudio sociodemográfico de la migración internacional. La reciente generalización de la difusión de los microdatos censales por parte de los institutos de estadística nacionales ha permitido a los microdatos de censos incorporarse al grupo de fuentes de información que tradicionalmente han nutrido al estudio de la migración internacional. Entre estas fuentes de información aparecen las encuestas que algunos países desarrollan de forma exclusiva para la investigación de la migración internacional, o los datos agregados extraídos de censos de población o de fuentes de flujos migratorios. Las limitaciones de ambas fuentes son bien conocidas, mientras que las encuestas no acostumbran a tener una cobertura total de los protagonistas de los flujos migratorios, los datos agregados carecen de información a nivel individual. Los microdatos censales también presentan sus limitaciones, pero logran superar las anteriormente citadas, y disponen de un gran potencial de análisis.

El contexto latinoamericano ha sido protagonista de muchos de los episodios recientes más importantes de movimientos migratorios transnacionales. La base de datos IPUMSI permite acercarse al estudio minucioso de muchas de estas etapas (Cuadro 3). La preservación de los microdatos censales de 1960 y 1970, especialmente en Argentina y Brasil, permite aproximarse a muchos de los individuos que protagonizaron las intensas relaciones migratorias que se establecieron durante la primera mitad de siglo XX entre numerosos países europeos, principalmente Italia y España, con América del Sur. En 1970 residían en Argentina más de un millón de personas nacidas en Italia o España, mientras que en el mismo tiempo residían en Brasil casi medio millón de portugueses y casi 200.000 japoneses.

Cuadro 3.- Población nacida en el extranjero y listado de los países de nacimiento con mayor presencia en los censos de América Latina disponibles en la base de datos IPUMS-International

	Nacidos en el extranjero	1 ^{er} país		2 ^o país		3 ^{er} país		4 ^o país		5 ^o país	
		Nombre	Población	Nombre	Población	Nombre	Población	Nombre	Población	Name	Población
Argentina 1970	2.207.000	Italia	668.850	España	541.600	Paraguay	233.750	Chile	145.150	Bolivia	103.350
Argentina 1980	1.921.498	Italia	488.068	España	379.360	Paraguay	265.432	Chile	218.488	Bolivia	119.332
Argentina 1991*	1.308.476	Italia	273.030	Paraguay	198.667	Chile	192.210	España	187.099	Uruguay	120.354
Argentina 2001	1.535.230	Paraguay	324.370	Bolivia	231.980	Italia	219.210	Chile	214.270	España	134.060
Brasil 1960	1.143.760	Portugal	293.980	Italia	203.620	Japón	157.120	España	148.520	Alemania	52.900
Brasil 1970	1.272.030	Portugal	457.860	Italia	160.850	Japón	157.945	España	136.375	Alemania	54.165
Brasil 1980	1.106.065	Portugal	392.205	Japón	137.325	Italia	106.830	España	100.430	Alemania	41.700
Brasil 1991	768.029	Portugal	262.958	Japón	82.990	Italia	67.035	España	56.613	Argentina	27.113
Brasil 2000	684.749	Portugal	212.928	Japón	69.717	Italia	55.305	España	42.976	Paraguay	29.517
Chile 1960	106.821										
Chile 1970	90.800	España	16.560	Argentina	13.570	Alemania	8.830	Bolivia	8.030	Italia	7.670
Chile 1982*	83.810	Argentina	19.640	España	12.250	Bolivia	6.620	Alemania	6.470	Italia	5.420
Chile 1992	115.030	Argentina	34.730	España	9.530	Perú	8.370	Bolivia	7.990	EE. UU.	6.080
Chile 2002	190.970	Argentina	50.330	Perú	37.760	Bolivia	11.690	Ecuador	10.060	España	9.630
Colombia 1964	71.050	Venezuela	16.050	Ecuador	10.050	EE. UU.	7.200	España	6.450	Alemania	3.900
Colombia 1973	72.870	Venezuela	20.420	Ecuador	9.300	EE. UU.	8.360	España	4.510	Alemania	2.780
Colombia 1985	78.396										
Colombia 1993	106.310	Venezuela	43.300	EE. UU.	13.940	Ecuador	9.260	España	3.550	Perú	3.150
Colombia 2005	104.600	Venezuela	37.025	EE. UU.	14.134	Ecuador	11.171	España	5.144	Perú	3.671
Costa Rica 1963	39.134	Nicaragua	20.230	Panamá	3.604	EE. UU.	2.074	Jamaica	1.989	España	1.394
Costa Rica 1973	47.220	Nicaragua	23.510	Panamá	4.290	EE. UU.	4.100	El Salvador	1.470	España	1.420
Costa Rica 1984	90.180	Nicaragua	45.970	El Salvador	8.880	EE. UU.	5.910	Panamá	5.090	Cuba	3.090
Costa Rica 2000	298.170	Nicaragua	227.770	Panamá	10.800	EE. UU.	9.150	El Salvador	9.060	Colombia	5.540

	Nacidos en el extranjero	1 ^{er} país		2 ^o país		3 ^{er} país		4 ^o país		5 ^o país	
		Nombre	Población	Nombre	Población	Nombre	Población	Nombre	Población	Nombre	Población
Ecuador 1962	27.522	Colombia	16.368	Perú	1.353	EE. UU.	1.089	España	858	Alemania	759
Ecuador 1974	44.607	*América	39.192	*Europa	4.523	*Asia	844	*África	36	*Oceanía	13
Ecuador 1982	76.340	*América	64.280	*Europa	9.640	*Asia	2.150	*Oceanía	160	*África	160
Ecuador 1990	77.980										
Ecuador 2001	104.870										
México 1960	153.564	EE. UU.	73.164	España	26.800	Guatemala	9.447	Alemania	3.618	Canadá	3.283
México 1970	189.600	EE. UU.	95.400	España	32.000	Guatemala	5.700	Alemania	5.700	Cuba	4.000
México 1990	339.220	EE. UU.	198.230	Guatemala	42.380	España	24.620	El Salvador	5.060	Colombia	4.660
México 1995	401.935	EE. UU.	284.961	*América	45.078	*Europa	43.811	*Asia	19.528	*África	400
México 2000	519.707	EE. UU.	358.399	Guatemala	29.156	España	21.334	Cuba	7.267	Canadá	7.245
Panamá 1960	82.640	Guatemala	40.620	Honduras	17.540	Nicaragua	4.340	Costa Rica	3.440	EE. UU.	3.320
Panamá 1970	51.110										
Panamá 1990	52.070										
Panamá 2000	81.690	Colombia	8.520	China	5.510	R. Dominicana	5.280	Nicaragua	5.280	EE. UU.	5.070
Venezuela 1971	583.616	Colombia	179.416	España	148.434	Italia	85.510	Portugal	58.198	EE. UU.	10.714
Venezuela 1981	1.045.470	Colombia	505.770	España	140.300	Portugal	92.100	Italia	79.280	Chile	25.110
Venezuela 1990	1.012.503	Colombia	532.740	España	102.888	Portugal	67.311	Italia	67.308	Perú	28.641
Venezuela 2000	1.008.900	Colombia	605.010	España	76.690	Portugal	53.270	Italia	48.170	Perú	34.040

Fuente: Elaboración propia a partir de los datos extraídos de <https://international.ipums.org>

* En los censos de Argentina 1991 y Chile 1982 hay una presencia muy destacada de individuos cuya respuesta es desconocida, por ello se produce el inesperado punto de inflexión en la evolución del total de población nacida en el extranjero.

Las muestras censales disponibles en la base de datos también permiten aproximarse a los movimientos migratorios transnacionales de carácter regional que se han producido en el interior de América Latina a lo largo de la segunda mitad de siglo XX, una vez frenadas las relaciones migratorias internacionales de larga distancia. Así, los microdatos de Venezuela, por ejemplo, revelan que en 1981 residían más de medio millón de personas naturales de Colombia. De la misma manera, también es posible reseguir la reciente intensificación de los flujos migratorios de nicaragüenses con destino a Costa Rica.

Pese a todo, es la presencia en la base de datos IPUMSI de los microdatos censales de los principales países hacia los que parten los flujos emigratorios latinoamericanos la que dota a esta fuente de información de un enorme valor para el análisis sociodemográfico de la migración internacional latinoamericana reciente (Cuadro 4).

El sistema IPUMSI incluye las muestras integradas de los censos de Estados Unidos, Canadá, España y Portugal. En el censo del año 2000 se contabilizaban en Estados Unidos casi quince millones de personas nacidas en países de América Latina, dos terceras partes de los cuales eran naturales de México. La fracción muestral de este censo es del 5%, lo que significa que hay más de 750.000 registros de individuos nacidos en América Latina y que residen en Estados Unidos puestos a la disposición del investigador. También es interesante el análisis de la emigración que se dirige a Canadá, donde residían en el año 2000 más de 600.000 personas nacidas en algún país de América Latina. No se puede obviar de este listado las relaciones migratorias con España, que se intensificaron durante los últimos años de la década de 1990. El censo de 2001 contabilizaba casi un millón de personas nacidas en América Latina, principalmente en Ecuador y Colombia.

Cuadro 4.- Población nacida en el extranjero y listado de los países latinoamericanos de nacimiento en los censos de los países del mundo con mayor presencia de población de América Latina (disponibles en IPUMSI)

	Nacidos en el extranjero	Nacidos en América Latina	1 ^{er} país		2 ^o país		3 ^{er} país		4 ^o país		5 ^o país	
			Nombre	Población	Nombre	Población	Nombre	Población	Nombre	Población	Nombre	Población
Estados Unidos 1960	10.789.197	819.238	México	582.539	Cuba	75.588	Argentina	16.735	Panamá	15.443	Brasil	14.253
Estados Unidos 1970	11.895.851	1.794.689	México	866.038	Cuba	449.909	Colombia	65.491	República Dominicana	62.704	Argentina	51.781
Estados Unidos 1980	16.241.594	3.949.787	México	2.243.655	Cuba	624.390	República Dominicana	168.344	Colombia	151.197	El Salvador	95.868
Estados Unidos 1990	22.906.830	7.654.216	México	4.409.033	Cuba	751.988	El Salvador	472.449	República Dominicana	353.755	Colombia	303.204
Estados Unidos 2000	34.656.270	14.737.463	México	9.325.452	Cuba	883.439	El Salvador	823.832	República Dominicana	698.106	Colombia	525.881
Estados Unidos 2005	39.331.908	17.615.403	México	11.164.770	El Salvador	994.418	Cuba	923.608	República Dominicana	729.244	Guatemala	652.909
Canadá 1971	3.299.800	102.600										
Canadá 1981	3.874.200	284.650										
Canadá 1991	4.637.663	492.900										
Canadá 2001	5.701.860	624.311										
España 1981	632.277	159.563	Argentina	34.869	Venezuela	30.680	Cuba	27.455	Brasil	12.449	Chile	10.839
España 1991*	835.788	107.880	Argentina	36.385	Venezuela	17.400	Resto de Sudamérica	41.131	Centroam.	8.230	Cuba	4.734

España 2001	2.147.880	834.540	Ecuador	220.240	Colombia	173.800	Argentina	104.040	Venezuela	68.420	Perú	52.620
Portugal 1981	394.820	27.480	Brasil	19.160	Venezuela	7.120						
Portugal 1991	458.120	42.180	Brasil	24.400	Venezuela	15.880						
Portugal 2001	642.720	73.940	Brasil	49.680	Venezuela	21.880						

Fuente: Elaboración propia a partir de los datos extraídos de <https://international.ipums.org>

*España 1991: El recuento de la población nacida en América Latina se ha tenido que elaborar a partir de la nacionalidad de la población

3.2.- Disponibilidad de variables para el estudio de la migración internacional en IPUMSI

La base de datos IPUMSI incluye en la actualidad alrededor de unas 13.000 variables, 500 de las cuales corresponden a variables integradas (Cuadro 5). Son éstas últimas las que, por tanto, han sido tratadas por el equipo IPUMSI con el objetivo de armonizar la información presente en los censos de población. Como es lógico, la calidad y cantidad de información que el sistema de extracción subministrará al investigador dependerá del contenido original del censo de población. Se puede hablar de cuatro grandes grupos de variables según su relación con el estudio de la migración internacional.

a) Variables de naturaleza y nacionalidad. En esta categoría se incluyen aquellas variables que permiten identificar el lugar de nacimiento y la nacionalidad de los individuos. Su disponibilidad varía según la variable y países elegidos.

Así, en el contexto de los países de América Latina, las variables de naturaleza y de lugar de nacimiento tienen una cobertura significativamente más amplia que la información referente a la nacionalidad. La naturaleza, que informa si el individuo ha nacido en el interior del país o en el extranjero, está disponible para todos los censos latinoamericanos que forman parte de la base de datos, con la única excepción del conteo de México 2005 y de Panamá 1980. Asimismo, se puede identificar el país de nacimiento en la gran mayoría de muestras. En cambio, las variables de nacionalidad tienen una cobertura significativamente menor, y únicamente están disponibles para las muestras de Argentina 1980, las 4 muestras de Costa Rica, México 1970 y Panamá 1970. En el caso de los microdatos de España y Estados Unidos, se puede identificar tanto el país de nacimiento como el país de nacionalidad en la mayor parte de las muestras de la base de datos.

En este mismo grupo de variables aparecen también aquellas que informan del año de llegada al país. Esta información está disponible para una tercera parte de las muestras latinoamericanas, aunque está más presente entre los censos más recientes: Argentina 1970 y 1980, Chile 1992 y 2002, Colombia 1973 y 2005, Costa Rica 1984 y 2000, Ecuador 2001 y Venezuela 1981 y 1991. Asimismo, el año de llegada al país está disponible en los últimos censos de España y Estados Unidos.

b) ***Variables de situación migratoria*** (Cuadro 6). Bajo esta agrupación aparecen las variables que revelan el itinerario migratorio de los individuos, normalmente relativo a los años más recientes a la fecha censal o al momento en que realizó el último cambio de residencia. Más de la mitad de las muestras censales latinoamericanas permiten conocer la situación migratoria con respecto al lugar en que el individuo residía 5 años atrás, o cuál es la relación geográfica del lugar de residencia actual en relación a la residencia anterior (la codificación de la variable, y las categorías disponibles se pueden consultar en el Cuadro 7). En algunos casos se puede conocer la duración de la residencia en la localidad actual, mientras que en las muestras de Colombia 2005 y México 2000 se puede conocer la razón que motivó el movimiento migratorio.

c) ***Variables construidas sobre interrelaciones familiares en el hogar***. Estas variables han sido construidas desde el propio proyecto IPUMSI y tienen como objetivo localizar a la madre, padre y pareja de todos los individuos, si es que los primeros residen en el mismo hogar. La creación de esta variable permite al sistema de extracción identificar la presencia de dichos familiares en el hogar, y asignar a los registros de cada individuo cualquiera de las características de los familiares. Esta aplicación puede ser de extraordinaria ayuda al investigador de la migración internacional ya que permite identificar de una forma sencilla a los hijos de padres migrantes así como la naturaleza de ambos miembros de la pareja.

Cuadro 7.- Ejemplo de codificación en el sistema IPUMSI de la variable “Situación migratoria, residencia anterior”

Código	Etiqueta del código	Brasil 1960	Brasil 1970	Brasil 1980	Brasil 1991	Brasil 2000	Chile 1960	Colombia 1973	Colombia 2005	Costa Rica 1963	Ecuador 1962	Ecuador 1974	México 1960	México 1970	México 1995	Panamá 1960	Panamá 1970	Panamá 1980	Panamá 2000	España 1991	España 2001
0	NIU (No en el Universo)	X	X	.	X	.
10	Misma unidad administrativa mayor	X	.	.	.	X	.	X	X	X	X
11	Misma unidad administrativa (mayor y menor)	X	X	X	X	.	X	X	X	.	X	X	X	X	X	.	X
12	Misma unidad administrativa mayor, diferente menor	X	X	X	X	.	X	X	X	.	X	X	X	X	X	X	X
20	Diferente unidad administrativa menor	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
30	En el extranjero	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	.	X	X	X
98	Respuesta eliminada	X	.	.
99	No declarado / Respuesta desconocida	.	X	.	X	X	X	.	X	X	X	X	X	.	X	X	X	X	X	.	.

Fuente: <https://international.ipums.org>

d) Resto de variables. Dependiendo de los intereses de la investigación, los usuarios pueden seleccionar cualquiera de las 500 variables integradas o 12.500 sin armonizar, estructuradas en el sistema a partir del tipo de registro (del hogar o individual). En el Cuadro 8 y 9 aparecen resumidos los grandes tipos de variables presentes en el sistema.

El proyecto IPUMSI planea incorporar a su base de datos la valiosa información correspondiente a los módulos sobre migración internacional de los censos de Colombia 2005, Ecuador 2001 y México 2000. Se prevé que dicha difusión se realice durante la primavera de 2009.

Cuadro 8.- Conjunto de variables del hogar disponibles en la base de datos IPUMSI

Variables del hogar (tal y como aparecen en el sistema de extracción de datos IPUMSI)	
Conjunto de variables	Detalle de las variables incluidas
Técnicas	Variables que permiten identificar el tipo de registro, la muestra, el número de familia dentro del hogar. Es fundamental la presencia del factor de ponderación de los registros de hogares.
Tipo de colectivo	Conjunto de variables que definen el tipo de colectivo que reside en el hogar: residencia familiar, establecimientos colectivos y vivienda vacante son las principales categorías.
Geografía	Variables geográficas que informan de la localización del hogar. En muchos de los países de América Latina se pueden encontrar hasta tres niveles administrativos.
Económicas	La principal variable de este grupo es el tipo de propiedad de la residencia, una variable disponible para la práctica totalidad de censos latinoamericanos.
Servicios	En esta categoría se incluyen las variables que hacen referencia a los servicios de los que dispone la vivienda: electricidad, agua corriente, aguas residuales, gas, calefacción, línea telefónica, conexión a Internet...
Electrodomésticos, vehículos y otros servicios	Conjunto de variables que permiten identificar la disponibilidad de aparatos electrónicos (tales como frigoríficos, televisores, aires acondicionados, ordenador personal...), automóviles y otros servicios en el hogar.
Características de la vivienda	Variables que definen el tipo de construcción, el material, el tamaño y la antigüedad de la vivienda.
Otras	En el caso de América Latina, aparecen en este grupo una variable que informa del número de personas fallecidas en el hogar durante el último año (Colombia 2005), así como la lengua principal del hogar (Ecuador 1990).
Variables construidas del hogar	Variables que han sido construidas por el equipo de IPUMSI, y que informan del número de familias en el hogar, del número de parejas casadas, así como del número de padres y madres en el hogar.

Fuente: <https://international.ipums.org>

Cuadro 9.- Conjunto de variables de la persona disponibles en la base de datos IPUMSI

Variables de persona (tal y como aparecen en el sistema de extracción de datos IPUMSI)	
Conjunto de variables	Detalle de las variables incluidas
Técnicas	Variables que permiten identificar el orden de la persona dentro del hogar. Es fundamental la presencia del factor de ponderación de los registros de personas.
Variables construidas sobre interrelaciones familiares	Variables construidas por el equipo IPUMSI, y que localizan a la madre, padre, y pareja de todos los individuos si es que los primeros residen en el hogar.
Puramente demográficas	Variables demográficas por excelencia. Entre ellas destacan la relación con la persona principal, la edad, el sexo, el estado civil, la edad al primer matrimonio, número de uniones, año y mes de nacimiento...
Fecundidad y mortalidad	En este conjunto de variables se incluye información relativa a la descendencia de los miembros del hogar. El número de hijos (nacidos o supervivientes) es una de las variables que aparecen en la mayoría de censos latinoamericanos.
Naturaleza y lugar de nacimiento	Variables que informan del lugar de nacimiento del individuo, así como de la nacionalidad. En algunas muestras también se puede conocer el año de llegada al país.
Etnicidad y lengua	Variables referentes a la raza, etnia y lengua/s hablada/s por el individuo. Los censos de Brasil y México interpelan también sobre la religión.
Educación	La asistencia a la escuela, el alfabetismo, el número de años en la educación, y el nivel de instrucción son variables que aparecen en la mayoría de censos. El nivel de instrucción se agrupa según una recodificación internacional, aunque se preserva también la codificación original del país.
Trabajo	En este apartado se incluyen las variables referentes al pasado y presente laboral del individuo. La relación con la actividad es una variable ampliamente recogida, así como el tipo de trabajador y la ocupación.
Ingreso	Algunas muestras de América Latina interpelan sobre los ingresos del individuo. Ése es el caso de Brasil, Costa Rica (1963), Panamá y Venezuela.
Migración	Aparecen aquí las variables que revelan el itinerario migratorio de los individuos, normalmente relativo a los años más recientes a la fecha censal o al momento en que realizó el último cambio de residencia.
Discapacidad	Es en este grupo donde aparecen variables que informan de la existencia de algún tipo de discapacidad del individuo. La introducción de estas cuestiones a los cuestionarios censales ha sido un hecho bastante reciente.
Otras	Aparecen aquí variables que no se han agrupado en ninguna de las categorías anteriores. En el caso de las muestras de América Latina tan sólo aparecen variables sobre el tipo de cobertura y asistencia sanitaria (México 2000).

Fuente: <https://international.ipums.org>

3.3.- Comparabilidad espacial y temporal de la información: nuevos enfoques en el estudio de la migración internacional

Tal y como se ha subrayado hasta el momento, una de las principales potencialidades de la base de datos IPUMSI radica en la comparabilidad espacial y temporal de las variables

integradas. En el caso preciso de la migración internacional, este elemento añade una serie de posibilidades analíticas con las que era difícil contar antaño. A continuación se recogen algunas líneas de investigación que la integración de los datos y la combinación de varias muestras censales permiten afrontar:

- a) Comparación de contextos migratorios de diferentes países. En este sentido, por ejemplo, sería posible contrastar la relación que se establece entre nivel de instrucción y la ocupación de la población de nacionalidad extranjera en dos campos inmigratorios diferentes, por ejemplo, Estados Unidos y España.
- b) Comparación de las características sociodemográficas y las trayectorias de los migrantes de un determinado país o ámbito en diferentes destinos. En este caso, por ejemplo, se podría analizar el comportamiento en términos de nupcialidad de la población ecuatoriana que emigra a España, la que se desplaza a Estados Unidos o la que se establece en otro país de América Latina.

Este tipo de perspectiva permite a IPUMSI sumarse, en cierta manera, a la iniciativa que desde hace años desarrolla CELADE, la base de datos IMILA, que permite reseguir las relaciones migratorias de los países latinoamericanos a través de los censos de población⁴.

- c) Comparación del contexto migratorio de un país a lo largo del tiempo. Sustrayendo información de todas las muestras censales de Estados Unidos, por ejemplo, se podría analizar si los inmigrantes latinoamericanos han cambiado su distribución en el territorio y su comportamiento en materia de movilidad residencial a lo largo de la segunda mitad de siglo XX.

Estas posibles líneas de investigación se añaden al propio interés que suscita *per se* el análisis de una única muestra censal y cuya tarea de extracción de datos también facilita el propio sistema IPUMSI.

⁴ <http://www.eclac.cl/migracion/imila>

4.- A modo de recapitulación

El estudio de la migración internacional ocupa un lugar principal entre los temas demográficos más tratados en América Latina. No en vano, según datos de 2005 residen en Estados Unidos casi 20 millones de personas nacidas en Latinoamérica, mientras que en España, según el avance del padrón de 2008, la cifra de nacidos en América Latina está a punto de alcanzar los dos millones. Pero no hay que olvidar, tampoco, la migración transnacional de carácter regional entre los países de América Latina, ni la larga tradición inmigratoria de la región.

En un momento en el que los microdatos censales se están consolidando como una de las principales fuentes de información para los investigadores en Ciencias Sociales, el proyecto IPUMS-International ha desarrollado una base de datos que contiene 111 muestras censales correspondientes a 35 países de todo el mundo. En ella se incluyen más de 250 millones de registros de personas que están a disposición de los investigadores a través de un sistema de extracción en línea que permite recibir la información deseada en unos minutos y sin coste económico alguno. Entre todos los países que integran la base de datos aparecen nueve de América Latina, más Estados Unidos y España, principales destinos de los flujos emigratorios recientes que parten de América Latina. La base de datos IPUMS constituye, así pues, una destacable fuente de información para el análisis de la migración internacional en el contexto latinoamericano.

Además de facilitar enormemente el propio proceso de recolección de datos censales, el sistema IPUMSI permite al investigador obtener la información de acuerdo a unos nuevos códigos integrados que hacen que dicha información sea comparable espacial y temporalmente. Este hecho abre la puerta a un amplio abanico de posibilidades de análisis en el estudio de la migración internacional. La falta de datos no parece ya una limitación para afrontar, por ejemplo, estudios comparativos en los que se combine información de diversas muestras censales.

Lamentablemente, no siempre todos los censos de población incorporan la información que el investigador requiere y que, quizás, otras muestras censales que se incluyen en la comparación sí que la proporcionan. Será ésta una de las limitaciones a las que el

investigador deberá hacer frente, así como al hecho de estar ante una integración compleja, ya que cada cuestionario censal responde a un diseño diferente.

Pese a todo, las principales características que permiten identificar a la población que ha traspasado sus fronteras están disponibles en la práctica totalidad de muestras de América Latina. También es importante destacar la aplicación del sistema que permite atribuir características del padre, madre o esposo al individuo, y así detectar la presencia de población extranjera a nivel familiar. Toda esta información se verá complementada en breve cuando se incluyan en la base de datos la información correspondiente a los módulos de migración internacional que se incluye en los censos más recientes de Colombia, Ecuador y México.

IPUMS-International es consciente del potencial que tienen los microdatos censales integrados en la investigación en Ciencias Sociales. Por ello trabaja activamente para poner a disposición de la comunidad científica series de microdatos integrados para el máximo número posible de países, en estrecha colaboración con los institutos de estadística nacionales, centros de investigación y profesionales de la Demografía. Entre los próximos retos del proyecto IPUMSI figura el de completar la integración de los censos de los países latinoamericanos y poder difundir sus datos. Para ello, se replicará la estrategia de distribución de datos de IPUMS-International con el objetivo de garantizar las máximas facilidades de accesos a los datos a nuestros usuarios.